63-3-2

Western Management Science Institute
University of California • Los Angeles

UNIVERSITY OF CALIFORNIA
Los Angeles

Western Management Science Institute

Working Paper No. 31

March 1963

A CHARACTERIZATION OF THE EXPONENTIAL DISTRIBUTION[1]

by

Thomas S. Ferguson

Supported by the Office of Naval Research under task number NR 047-041 and the Western Management Science Institute under its grant from the Ford Foundation.

A CHARACTERIZATION OF THE EXPONENTIAL DISTRIBUTION[1]

by Thomas S. Ferguson

University of California, Los Angeles

1. *Introduction and summary*. This paper continues a study, initiated in [1], of the stochastic independence of $V = X - Y$ and the mean of order γ,

$$(1) \qquad M_\gamma = \begin{cases} (1/\gamma) \log [(e^{\gamma X} + e^{\gamma X})/2] & \text{for } \gamma \neq 0 \text{ and } \gamma \neq \pm\infty \\ (X + Y)/2 & \text{for } \gamma = 0 \\ \max (X,Y) & \text{for } \gamma = +\infty \\ \min (X,Y) & \text{for } \gamma = -\infty \end{cases}$$

for independent X and Y.

The case $\gamma = 0$ is solved by the well known result, *if, for independent* X *and* Y, *the variables* X + Y *and* X - Y *are independent, then both* X *and* Y *have normal distributions with a common variance.* The first proof of this result may be found in Kac [5]. (Although the statement of this result there is much weaker, Kac's proof applies exactly to the statement above.)

A corresponding result for the gamma distribution, proved in restricted form by Hogg [4], and proved without restrictions by Lukacs [6], is as follows: *if* Z_1 *and* Z_2 *are independent positive random variables for which* $Z_1 + Z_2$ *and* Z_1/Z_2 *are independent, then both* Z_1 *and* Z_2 *have gamma distributions with a common scale parameter.* This

result may be used to complete the study of the independence of M_γ and V for independent X and Y when γ is finite and non-zero, by letting $Z_1 = \exp(\gamma X)$ and $Z_2 = \exp(\gamma Y)$. It is clear, then, from the result of Hogg and Lukacs that both $\exp(\gamma X)$ and $\exp(\gamma Y)$ must have gamma distributions with a common scale parameter. The resulting distributions of X and Y have been studied in [1].

Therefore, of this study, there remain yet to be completely solved only the cases $\gamma = +\infty$ and $\gamma = -\infty$. These two cases are essentially the same since if $M_{+\infty}$ and V are independent, then so are $\min(-X,-Y)$ and V. Thus, the negatives of the distributions for which $M_{+\infty}$ and V are independent will yield independent $M_{-\infty}$ and V. We consider only the case $\gamma = -\infty$. This problem is considered in a previous paper, [2], in which the distributions were restricted to be discrete. The main result of that paper is as follows: _if_ X _and_ Y _are independent_, _non-degenerate_, _discrete random variables_, _and if_ $U = \min(X,Y)$ _and_ $V = X - Y$ _are independent_, _then_ X _and_ Y _both have geometric distributions with common location and scale parameters_, _but possibly different geometric parameters_. If the probability mass function of the geometric distribution is written as

$$p(x) = (1 - p)\, p^{(x - \theta)/c} \qquad \text{for } x = \theta, \theta + c, \theta + 2c, \cdots \tag{2}$$

then, the parameter θ is a location parameter, c is a scale parameter, and p is called the geometric parameter, $c > 0$, $0 < p < 1$.

In this paper, the distributions of X and Y are restricted to be absolutely continuous. This will result in a characterization of the so-called exponential distribution whose density is

(3) $$f(x) = (1/\sigma) \exp(-(x - \Theta)/\sigma) \qquad \text{if } x > \Theta$$

and zero otherwise. The parameter Θ is a location parameter, and the parameter $\sigma > 0$ is a scale parameter, which is also the standard deviation of the distribution. In these terms the main result of this paper may be stated as follows: *if X and Y are independent random variables with absolutely continuous distributions, and if* $U = \min(X,Y)$ *and* $V = X - Y$ *are independent, then both X and Y have exponential distributions with a common location parameter but with possibly different scale parameters*. This result then gives a characterization of the exponential distribution, since it is easy to show the converse, that if X and Y are independent and if each has an exponential distribution with a common location parameter but with possibly different scale parameters, then U and V are independent.

2. A characterization of the exponential distribution. The following three hypotheses are referred to frequently in this section.

H_1: *The random variables X and Y are independent.*

H_2: *The random variables* $U = \min(X,Y)$ *and* $V = X - Y$ *are independent.*

H_3: *The distributions of X and Y are absolutely continuous (with respect to Lebesgue measure).*

Hypothesis H_3 merely states that the probability densities (with respect to Lebesgue measure), $f_X(x)$ and $f_Y(y)$, of the variables X and Y exist. It is immediate from these hypotheses that the distributions of U and V are absolutely continuous also. Furthermore, the joint distribution of U and V has a density which factors from hypothesis H_2,

(4) $$f_{U,V}(u,v) = f_U(u)\ f_V(v)$$

and, using hypothesis H_1 also, is easily found to be

(5) $$f_U(u)\ f_V(v) = f_X(u + v)\ f_Y(u) \qquad \text{for } v > 0$$

and

(6) $$f_U(u)\ f_V(v) = f_X(u)\ f_Y(u - v) \qquad \text{for } v < 0$$

except perhaps for (u,v) _in some null set_, N, _in the plane_. In the following, Lebesgue measure on the real line will be denoted by ℓ, and Lebesgue measure on the plane will be denoted by ℓ_2. Thus, equations (5) and (6) are valid unless $(u,v) \in N$, where $\ell_2(N) = 0$.

It is our desire to show that, under hypotheses H_1 H_2 and H_3, both X and Y have exponential distributions with a common location parameter, but with possibly different scale parameters. In the demonstration of this fact, we shall work with equations (5) and (6), and as a consequence we shall have to pay a great deal of attention to null sets, thus obscuring the proof. However, a general outline of the proof may be found in the statements of the several lemmas which preceed the main theorem.

Lemma 1. _If_, _for independent non-degenerate_ X _and_ Y, U _and_ V _are independent_, _then_ $P(V > 0) > 0$ _and_ $P(V < 0) > 0$.

Proof. Suppose $P(V > 0) = 0$; then,

(7) $$P(X = U) = P(X \leq Y) = P(V \leq 0) = 1.$$

Thus, the statement that U and V are independent is the same as the statement that X and X - Y are independent. But from lemma 1 of [2], X cannot be independent of X - Y for independent X and Y unless X is degenerate. This contradiction proves $P(V > 0) > 0$. Consideration of -V in place of V proves $P(V < 0) > 0$ by symmetry.

<u>Lemma 2</u>. <u>If</u>, <u>for independent non-degenerate</u> X <u>and</u> Y, U <u>and</u> V <u>are</u> <u>independent</u>, <u>then the distributions of</u> U, X, <u>and</u> Y <u>are not bounded</u> <u>above</u>.

<u>Proof</u>. We shall prove the lemma for the distribution of U, since the result for the distributions of X and Y obviously follow from this. Suppose the distribution of U is bounded above, and let b be the least upper bound for the distribution of U. Thus

$$P(U > b) = 0 \tag{8}$$

and, for every number $\epsilon > 0$,

$$P(U > b - \epsilon) > 0 \tag{9}$$

Now, find an $\epsilon > 0$ such that $P(V > \epsilon) > 0$. That such an ϵ exists follows from lemma 1. Then,

$$\begin{aligned} 0 &< P(V > \epsilon)\ P(b - \epsilon < U \leqq b) \\ &= P(V > \epsilon,\ b - \epsilon < U \leqq b) \\ &= P(X > Y + \epsilon,\ b - \epsilon < Y \leqq b) \\ &\leqq P(X > b)\ P(b - \epsilon < Y \leqq b)\ . \end{aligned} \tag{10}$$

Hence, $P(X > b) > 0$. But by symmetry, $P(Y > b) > 0$, which implies, by the definition of U, that $P(U > b) > 0$, contrdicting (8) and proving the lemma.

Let A_U represent the null set of the density of U; i.e. $A_U = \{u: f_U(u) = 0\}$. Similarly, let $A_X = \{u: f_X(u) = 0\}$ and $A_Y = \{u: f_Y(u) = 0\}$. The following lemma shows that these three sets are essentially identical. The notation $E_1 \Delta E_2$ represnets the symmetric difference of the sets E_1 and E_2; i.e. $E_1 \Delta E_2 = E_1E_2^c \cup E_2E_1^c$. We note that H_3 implies that both X and Y are non-degenerate, so that lemmas 1 and 2 are valid under hypotheses H_1 H_2 and H_3.

Lemma 3. Under hypotheses H_1 H_2 and H_3, $\ell(A_X \Delta A_Y) = \ell(A_X \Delta A_U) = \ell(A_Y \Delta A_U) = 0$.

Proof. Let $u_0 \in A_Y$; then, for each $v > 0$, equation (5) implies that $f_U(u_0)\, f_V(v) = 0$ unless $(u_0,v) \in N$. But lemma 1 implies that $f_V(v) > 0$ on a set, call it B, of positive Lebesgue measure in the interval $(0, \infty)$. Thus, if $f_U(u_0) \neq 0$, then $(u_0,v) \in N$ for all $v \in B$. This implies that if $u_0 \in A_Y$, then $u_0 \in A_U$ unless $u_0 \in N^* = \{u: (u,v) \in N$ for all $v \in B\}$; or, more simply,

$$A_Y \subset A_U \cup N^* . \tag{11}$$

Furthermore, $\ell(N^*) = 0$, since

$$\ell(N^*)\, \ell(B) = \ell_2(N^* \times B) \leq \ell_2(N) = 0 . \tag{12}$$

Now suppose that $u_0 \in A_U$; then, for each $v > 0$, equations (5) implies that $f_X(u_0 + v)\, f_Y(u_0) = 0$, unless $(u_0, v) \in N$. But lemma 2 implies that $f_X(u) > 0$ on a set B_{u_0} of positive measure in the interval (u_0, ∞). Thus, if $f_Y(u_0) \neq 0$, then $(u_0, v) \in N$ for all $v \in B_{u_0} - u_0$. Hence,

$$A_U \subset A_Y \cup \hat{N}, \tag{13}$$

where $\hat{N} = \{u\colon (u,v) \in N \text{ for all } v \in B_u - u\}$. If we let $E = \{(u,v)\colon u \in \hat{N} \text{ and } v + u \in B_u\}$, so that $E \subset N$ and $\ell_2(E) = 0$, then, by Fubini's theorem, of more directly by theorem A page 147 of Halmos [3], almost every section E_u has measure zero. But for $u \in \hat{N}$, E_u has positive measure, implying that $\ell(\hat{N}) = 0$.

Together, equations (11) and (13) imply

$$\ell(A_Y \,\Delta\, A_U) = \ell(N^*) + \ell(\hat{N}) = 0. \tag{14}$$

By symmetry, $\ell(A_X \,\Delta\, A_U) = 0$, and as a consequence, $\ell(A_X \,\Delta\, A_Y) = 0$, completing the proof.

Lemma 3 implies that the sets on which f_X, f_Y, and f_U vanish may be taken to coincide by increasing if necessary, the null set N outside of which equations (5) and (6) hold. This we shall assume done, and we shall denote the set on which f_X (or f_Y or f_U) vanishes by A, and we shall assume that f_X (and f_Y and f_U) is positive on A^c.

We may choose any representations of the densities we like for these formulas; all such representations differ only on null sets. However, it will be convenient in what follows to choose as the particular representative of the density of V

$$(15) \qquad f_V(v) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(x - v)\, dx$$

for all v, for any choice of the densities of X and Y. It will turn out that this function is already a continuous function of v over the whole real line.

<u>Lemma 4</u>. <u>Under hypotheses</u> H_1 H_2 <u>and</u> H_3, <u>there exist an</u> $\epsilon > 0$ <u>and an</u> $\eta > 0$ <u>such that</u> $f_V(v) \geq \eta$ <u>for all</u> v <u>such that</u> $-\epsilon \leq v \leq \epsilon$.

<u>Proof</u>. Let $D_n = \{u: f_X(u) > 1/n\}$ and $E_n = \{u: f_Y(u) > 1/n\}$. Then both $D_n \nearrow A^c$ and $E_n \nearrow A^c$. Furthermore, $D_n \cap E_n \nearrow A^c$, and using (15)

$$(16) \qquad f_V(v) \geq \int_{D_n \cap (E_n - v)} f_X(x)\, f_Y(x - v)\, dx$$

$$\geq (1/n^2)\, \ell(D_n \cap (E_n - v)) \ .$$

First we find an n such that

$$(17) \qquad \ell(D_n \cap E_n) > \min(1, \ell(A^c)/2) \ .$$

(Note that $\ell(A^c)$ may be infinite.) Then, since (exercise (1) page 268 of Halmos [4])

$$(18) \qquad \ell(D_n \cap (E_n - v)) \to \ell(D_n \cap E_n) \ , \qquad \text{as } v \to 0,$$

there exists an $\epsilon > 0$ such that for $-\epsilon \leq v \leq \epsilon$,

$$(19) \qquad \ell(D_n \cap (E_n - v)) \geq \min(1, \ell(D_n \cap E_n)/2) \ .$$

This implies that for $-\epsilon \leq v \leq \epsilon$,

(20) $$f_V(v) \geq (1/n^2)\min(1/2, \ell(A^c)/2) = \eta > 0 ,$$

proving the lemma.

The following lemma shows that the null set, A, of the distribution of X, is essentially the interval $(-\infty,\theta)$, where θ is some number ($\theta = -\infty$ being a possibility).

Lemma 5. Under hypotheses H_1 H_2 and H_3, either (i) $\ell(A) = 0$ or (ii) there exists a finite number θ such that $\ell(A\,\Delta(-\infty,\theta)) = 0$.

Proof. We first show that for almost all $u \in A^c$, $\ell((A - u) \cap (0,\epsilon)) = 0$, where ϵ is any positive number such that $f_V(v) > 0$ for $0 < v < \epsilon$. Such an ϵ exists from lemma 4. Since $\ell((A - u) \cap (0,\epsilon))$ is a continuous function of u, (again Halmos [3] page 268), the set $C_1 = \{u: \ell((A - u) \cap (0,\epsilon)) > 0\}$ is open. Thus the set $C_2 = A^c \cap C_1$ is measurable, and we are to show $\ell(C_2) = 0$. If $u \in C_2$, then $(u,v) \in N$ for every v for which $0 < v < \epsilon$ and $u + v \in A$, since for such a v equation (5) is not valid. Now let $D = \{(u,v): u \in C_2, (u,v) \in N\}$ so that $D \subset N$. Then every section of D, $D_u = \{v: (u,v) \in D\}$, for which $u \in C_2$, has positive Lebesgue measure. Then, again by Halmos [3] theorem A page 147, $\ell(C_2) = 0$ since $\ell_2(D) = 0$.

Thus for almost all $u \in A^c$, $\ell((A - u) \cap (0,\epsilon)) = 0$. By induction, $\ell((A - u) \cap (0,(3/2)^n\epsilon)) = 0$ for all n and for almost all $u \in A^c$. Since true for all n, we have

(21) $$\ell(A \cap (u,\infty)) = 0 \qquad \text{for almost all } u \in A^c$$

Either (i) $\ell(A) = 0$ or (ii) $\theta = \text{glb}\{u: \ell(A \cap (u,\infty)) = 0\}$ exists. In the second case, it is easy to see from (21) that $\ell(A\Delta(-\infty,\theta)) = 0$ as was to be proved.

According to this lemma, we may take the set A to be the set $(-\infty,\theta)$ (if A is a null set, $\theta = -\infty$) by enlarging, if necessary, the set N. We assume that this has been done. Thus each of the functions f_X, f_Y, and f_Z are zero of the set $(-\infty,\theta)$ and positive of the set (θ,∞).

Lemma 6. Under hypotheses H_1 H_2 and H_3, $f_V(v)$ is equal almost everywhere to a continuous, infinitely differentiable and positive function of v on the intervals $(-\infty,0)$ and $(0,\infty)$.

Proof. For $u > \theta$, equation (5) asserts that

$$f_X(u + v) = f_V(v)\ [f_U(u)/f_Y(u)] \tag{22}$$

for $v > 0$ and for $(u,v) \notin N$. Let $g(u)$ be any function of u for which

a) $g(u) = 0$ outside (b_1,b_2) where $\theta < b_1 < b_2$,

b) $g(u) > 0$ inside (b_1,b_2), and

c) $g(u)$ is continuous and has n continuous derivatives on the real line, where n is an arbitrary positive integer. Then

$$\int f_X(u + v)\ g(u)\ du = f_V(v) \int [f_U(u)/f_Y(u)]\ g(u)\ du \tag{23}$$

for almost all $v > 0$. Since the left side is finite and positive for all $v > 0$, the integral on the right side must be finite and positive also. But the left side is equal to

$$\int f_X(w)\ g(w - v)\ dw \tag{24}$$

which is a continuous function of v with n continuous derivatives. Since $f_V(v)$ is equal for almost all $v > 0$ to (24) divided by the integral on the right side of (23), and since n is arbitrary, the lemma is proved for the interval $(0,\infty)$. By symmetry, the lemma must also be true for the interval $(-\infty,0)$.

<u>Lemma 7</u>. <u>Under hypotheses</u> H_1 H_2 <u>and</u> H_3, <u>each of the functions</u> $f_X(x)$, $f_Y(y)$, <u>and</u> $f_U(u)$, <u>is equal almost everywhere in the interval</u> (θ,∞) <u>to a continuous</u>, <u>infinitely differentiable function</u>.

<u>Proof</u>. From (22), it is immediate that $f_X(u)$ is equal for almost all $u > \theta$ to a continuous, infinitely differentiable function. That $f_Y(u)$ is also equal almost everywhere on the interval (θ,∞) to such a function follows from symmetry. Then, dividing equation (5) by $f_V(v)$ (we may assume it is positive for $v > 0$ by enlarging N if necessary) will immediately give the same result for $f_U(u)$, completing the proof.

By enlarging the set N again if necessary, we shall choose $f_X(x)$, $f_Y(y)$, and $f_U(u)$ to be continuous and infinitely differentiable on the interval (θ,∞) and to vanish on the interval $(-\infty,\theta)$. The fact that the four functions, f_X, f_Y, f_U, and f_V, are continuous except possibly at one point, implies that equation (5) must be valid for all $u \neq \theta$ and for all $v > 0$. Similarly, equation (6) must be valid for all $u \neq \theta$ and for all $v < 0$.

<u>Theorem</u>. <u>Suppose that the random variables</u> X <u>and</u> Y <u>are independent and have absolutely continuous distributions</u>. <u>Then</u>, <u>in order that</u> $U = \min(X,Y)$ <u>and</u> $V = X - Y$ <u>be independent</u>, <u>it is necessary and sufficient that both</u> X <u>and</u> Y <u>have exponential distributions with the same value of the location parameter</u>.

<u>Proof</u>. <u>Necessity</u>: Since equation (5) is now valid and positive for all $u > \theta$ and for all $v > 0$, we may take logarithms of both sides, and differentiate both sides first with respect to $v > 0$ and then with respect to $u > \theta$. Denoting the logarithm of f_X by h_X, we have

$$h_X''(u) = 0 \tag{25}$$

for all $u > \theta$. The solution of this differential equation gives for $u > \theta$

$$\log f_X(u) = \alpha u + \beta \tag{26}$$

where α and β are constants of integration. Since $f_X(u)$ vanishes for $u < \theta$, and since the area under the curve $f_X(u)$ must be finite, it is clear that α is negative and θ is finite. Let $\sigma_1 = -\alpha^{-1}$. Then, since the area under the curve $f_X(u)$ must be equal to one, it follows that $-\beta = \theta/\sigma_1 + \log \sigma_1$. Hence,

$$f_X(x) = (1/\sigma_1) \exp(-(x - \theta)/\sigma_1) \qquad \text{if } x > \theta \tag{27}$$

and zero otherwise. By symmetry.

$$f_Y(y) = (1/\sigma_2) \exp(-(y - \theta)/\sigma_2) \qquad \text{if } y > \theta \tag{28}$$

and zero otherwise (that the same value of θ must be used follows from lemma 3).

<u>Sufficiency</u>. Now suppose that the distributions of X and Y are given by (27) and (28). Then the joint density of U and V may be written

$$f_{U,V}(u,v) = \begin{cases} f_X(u + v)\, f_Y(u) & \text{for } v > 0 \\ \\ f_X(u)\, f_Y(u - v) & \text{for } v < 0 \end{cases}$$

$$(29) \qquad = (1/\sigma_1\sigma_2) \begin{cases} 0 & \text{for } u < \theta \\ \exp[-(\frac{1}{\sigma_1} + \frac{1}{\sigma_2})(u - \theta) - \frac{v}{\sigma_1}] & \text{for } u > \theta,\ v > 0 \\ \exp[-(\frac{1}{\sigma_1} + \frac{1}{\sigma_2})(u - \theta) + \frac{v}{\sigma_2}] & \text{for } u > \theta,\ v < 0 \end{cases}$$

$$= f_U(u)\ f_V(v),$$

where

$$(30) \qquad f_U(u) = \begin{cases} 0 & \text{for } u < \theta \\ (\frac{1}{\sigma_1} + \frac{1}{\sigma_2}) \exp(-(\frac{1}{\sigma_1} + \frac{1}{\sigma_2})(u - \theta)) & \text{for } u > \theta \end{cases}$$

and

$$(31) \qquad f_V(v) = \frac{1}{\sigma_1 + \sigma_2} \begin{cases} \exp(-v/\sigma_1) & \text{for } v > 0 \\ \exp(+v/\sigma_2) & \text{for } v < 0 \end{cases}$$

proving the independence of U and V.

We note from equations (30) and (31) that U has an exponential distribution while V has a double exponential distribution with different scale parameters for the two sides and with the weights chosen so that the density is continuous at the origin.

3. Concluding remarks. Here it will be pointed out what remains to be solved of the problem, mentioned in the introduction, of finding distributions for independent random variables, X and Y, which make

$U = \min(X,Y)$ and $V = X - Y$ independent. This paper solves the problem when both X and Y have absolutely continuous distributions. A preceeding paper, [2], solves the problem when both X and Y are discrete, or when one of X or Y is degenerate. Still unsettled are the cases in which at least one of the distributions is continuous singular, or a non-trivial mixture of an absolutely continuous, discrete, and continuous singular distribution. In the following paragraph, an argument will be given which will dispose of the case in which at least one of the distributions has a positive probability mass at some point. What remains, therefore, is the problem of finding whether or not there are continuous distributions for independent X and Y, at least one of these distributions not absolutely continuous, for which U and V are independent.

Now suppose that X and Y are independent, non-degenerate random variables such that $U = \min(X,Y)$ and $V = X - Y$ are independent. We suppose further that either the distribution of X or the distribution of Y has at least one point mass. We will show that the distributions of both X and Y must be discrete, so that from the main theorem of [2], both X and Y must have geometric distributions with common location and scale parameters. Suppose there exists a number, x_0, such that $P(X = x_0) > 0$. Then from lemma 2 of this paper, $P(Y > x_0) > 0$, so that from lemma 2 of [2], $P(Y = x_0) > 0$. Hence, $P(V = 0) > 0$. But since U and V are independent,

$$
\begin{aligned}
P(X = Y \leq u) &= P(U \leq u,\ V = 0) \\
&= P(U \leq u)\ P(V = 0) \qquad (32) \\
&= (1 - P(X > u)\ P(Y > u))\ P(V = 0).
\end{aligned}
$$

Hence, any time there is an increase on the right side as u increases, there is a corresponding increase on the left. That is to say, any interval containing some mass of the distribution of either X or Y contains a common point mass of each. Non-degeneracy implies there must be at least two of these, and the proof of theorem 1, [2], implies that all such mass points must lie on a fixed lattice. Together, these facts imply that the distributions of both X and Y must be discrete.

References

[1] Ferguson, T. S. (1962) Location and scale parameters in exponential families of distributions. Ann. Math. Statist. **33** 986-1001.

[2] Ferguson, T. S. (to appear) A characterization of the geometric distribution. Submitted to Ann. Math. Statist.

[3] Halmos, P. R. (1950) Measure Theory Van Nostrand.

[4] Hogg, R. V. (1951) On ratios of certain algebraic forms Ann. Math. Statist. **22** 567-572.

[5] Kac, M. (1959) On a characterization of the normal distribution Amer. Jour. Math. **61** 726-728.

[6] Lukacs, E. (1955) A characterization of the gamma distribution Ann. Math. Statist. **26** 319-324.

[1]This work was supported partly by the Office of Naval Research under task 047-041 and partly by the Western Management Science Institute under a grant from the Ford Foundation.